Lecture Notes on MAT3040 Advanced Linear Algebra

by Aaron Laai

The Chinese University Of Hong Kong, Shenzhen
March 19, 2019

Preface

This book will mainly follow the path induced in the course MAT3040 Advanced Linear Algebra. It was lectured by Prof. Daniel Wong in 2019 spring. The major materials covered in class will be contained. However, this book is written for reviewing rather than self-learning. For this reason, and for the sake of saving time, some modifications will be made to keep the content neat, concise, and suitable for grasping the outline of the course. In addition, compared to those in class, some theorems will be presented in different manners, some proofs will follow different methods, and some remarks will be given in paragraphs to make the logic clear.

Table of contents

1 Preliminaries

Some differences between 2040 and 3040:

- $\mathbb{R}^n$ vs general vector space V
- Linear transformation $T\cdot\mathbb{R}^n\to\mathbb{R}^n$ vs $T\cdot V\to W$
- Eigenvalues of $A_{n\times n}\to$ Eigenvalues of a linear operator T
- Dot product $\to$ Inner product $<\boldsymbol{v}_1,\boldsymbol{v}_2>$

Example. $C(\mathbb{R})$, the space of all continuous functions in $\mathbb{R}$. $C^\infty(\mathbb{R})$. $\mathbb{R}[x]$

Example. Laplace equation:

$$\nabla f=0$$

is a linear operator with

$$\begin{aligned}\nabla\colon C^\infty(\mathbb{R}) &\mapsto C^\infty(\mathbb{R})\\ f &\mapsto \left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)f\end{aligned}$$

1 Vector Spaces

Definition 1.1. *A vector space V over a field $\mathbb{F}$ is a set closed under addition and scalar multiplication of objects in it subject to the following rules*

i. *(Commutativity of vector addition)* $\forall v_1,v_2\in V, v_1+v_2=v_2+v_1$.
ii. *(Associativity of vector addition)* $v_1+(v_2+v_3)=(v_1+v_2)+v_3$.
iii. *(Additive identity)* $\exists!0\in V$ *such that* $0+v=v,\forall v\in V$.
iv. *(Additive inverse)* $\forall v\in V,\exists -v\in V$ such that $v+(-v)=(-v)+v=0$.
v. *(Associativity of scalar multiplication)* $\alpha_1(\alpha_2 v)=(\alpha_1\alpha_2)v$.
vi. *(Commutativity of scalar)* $(\alpha_1\alpha_2)v=(\alpha_2\alpha_1)v$.
vii. *(Multiplicative identity)* $1\cdot v=v$.
viii. *(Distribution)*
   - $(\alpha_1+\alpha_2)v=\alpha_1 v+\alpha_2 v$, $\alpha(v_1+v_2)=\alpha v_1+\alpha v_2$
   - $0\cdot v=0$

In algebraic sense we can say that a vector space V equipped with a scalar field $\mathbb{F}$ is an abelian group under vector addition, and that has associative and commutative scalar multiplication subject to distribution laws. Besides, any vector is fixed when multiplied with the multiplicative identity of $\mathbb{F}$.

In the following text, unless specified otherwise, we assume the scalar field to be $\mathbb{F}$.

Example 1.2. Apply all these rules when $V = \mathbb{F}^n, M_{n\times n}(\mathbb{F}), \text{and} \, C(\mathbb{R})$.

Definition 1.3. *A subset $W \subseteq V$ of a vector space V is called a vector subspace of V if W itself forms a vector space, denoted by $W \leqslant V$.*

Example 1.4. Let $V = \mathbb{R}^3$. Let W be a plane containing 0 in $\mathbb{R}^3$, i.e., $W = \{(x, y, 0)|x, y \in \mathbb{R}\}$. Then $W \in V$(check this). Let $W' = \{(x, y, 1)|x, y \in \mathbb{R}\}$. Then W' is not a subspace of V since W' is not closed under addition.

Proposition 1.5. *$W \subseteq V$ is a subspace of V if and only if $\forall w_1, w_2 \in W$, $\alpha_1, \alpha_2 \in \mathbb{F}$, $\alpha_1 w_1 + \alpha_2 w_2 \in W$.*

Example 1.6. $V = M_{n\times n}(\mathbb{F})$. $W = \{A \in V | A^T = A\}$ is a subspace of V.

Example 1.7. $V = C^\infty(\mathbb{R})$. $W = \left\{ f \in V | \dfrac{d^2}{dx} f + f = 0 \right\}$ is subspace of V. It is clear since $\dfrac{d^2}{dx}$ is a linear operator.

Definition 1.8. *Let V be a vector space over $\mathbb{F}$.*

i. A linear combination of $S \subseteq V$ is of the form $\sum_{i=1}^n \alpha_i s_i$, $\alpha_i \in \mathbb{F}$, $s_i \in S$. (Finite sum)

ii. The span of a subset S of V is $\mathrm{Span}(S) = \{\sum_{i=1}^n \alpha_i s_i | \alpha_i \in \mathbb{F}, s_i \in S\}$

iii. S is a spanning set of V or, S spans V, if $V = \mathrm{Span}(S)$.

Example 1.9. $V = \mathbb{R}[x]. S = \{1, x^2, x^4, \ldots\}$. Then $2 + x^4 + \pi x^{106} \in \mathrm{Span}(S)$ but $1 + x^2 + x^4 + \cdots \notin \mathrm{Span}(S)$. Obviously, $\mathrm{Span}(S) \neq V$. However, S is the spanning set of $W = \{p \in V | p(x) = p(-x)\}$.

Example 1.10. $V = M_{3\times 3}(\mathbb{R}). W_1 = \{A \in V | A^T = A\}. W_2 = \{B \in V | B^T = -B\}$(Skew symmetric). Both of them are subspace of V. Let $S := W_1 \cup W_2$. It is an exercise to show that S spans V.

Lemma 1.11. *Let S be a subset in vector space V. Then*

i. $S \subseteq \mathrm{Span}(S)$

ii. $\mathrm{Span}(S) = \mathrm{Span}(\mathrm{Span}(S))$

iii. If $w \in \mathrm{Span}\{v_1, v_2 \ldots, v_n\} \backslash \mathrm{Span}\{v_2 \ldots, v_n\}$, then $v_1 \in \mathrm{Span}\{w, v_2 \ldots, v_n\} \backslash \mathrm{Span}\{v_2 \ldots, v_n\}$.

Proof.

i. For $s \in S$, $s = 1 \cdot s \in \mathrm{Span}(S)$

ii. By i, $\mathrm{Span}(S) \subseteq \mathrm{Span}(\mathrm{Span}(S))$. Pick

$$v = \sum_{i=1}^n \alpha_i v_i \in \mathrm{Span}(\mathrm{Span}(S))$$

where $v_i \in \mathrm{Span}(S)$. Write $v_i = \sum_j^{m_i} \beta_{\mathrm{ij}} s_j$ where $s_j \in S$. Then

$$\begin{aligned} v &= \sum_{i=1}^{n} \alpha_i v_i \\ &= \sum_{i=1}^{n} \alpha_i \left(\sum_j^{m_i} \beta_{\mathrm{ij}} s_j \right) \\ &= \sum_j^{m_i} \left(\sum_{i=1}^{n} (\alpha_i \beta_{\mathrm{ij}}) \right) s_j \\ &\in \mathrm{Span}(S) \end{aligned}$$

iii. By hypothesis, $w = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ is not the zero vector since the both $\mathrm{Span}\{v_1, v_2 ..., v_n\}$ and $\mathrm{Span}\{v_2 ..., v_n\}$ contain the zero vector. Without loss of generality (abbreviated to WLOG in the following text), assume that $\alpha_1 \neq 0$. Then

$$v_1 = \left(-\frac{\alpha_2}{\alpha_1} \right) v_2 + \cdots + \left(-\frac{\alpha_n}{\alpha_1} \right) v_n + \left(\frac{1}{\alpha_1} \right) w$$

Thus $v_1 \in \mathrm{Span}\{w, v_2 ..., v_n\}$. Suppose on the contrary that $v_1 \in \mathrm{Span}\{v_2 ..., v_n\}$. We claim that $\mathrm{Span}\{v_1, v_2 ..., v_n\} = \mathrm{Span}\{v_2 ..., v_n\}$(left as exercise). But then

$$w \in \phi = \mathrm{Span}\{v_1, v_2 ..., v_n\} \backslash \mathrm{Span}\{v_2 ..., v_n\}$$

with ϕ being a empty set, which is a contradiction. The proof completes here.

□

2 Linear Independence and Basis

Definition 2.1. *Let S be a subset (not necessarily finite) of V. Then S is linearly independent on V if for any finite subset $\{s_1, ..., s_k\}$ in S, $\sum_{i=1}^{k} \alpha_i s_i = 0$ if and only if $\alpha_i = 0, \forall i$. $(\alpha_1, \alpha_2, ..., \alpha_k)$, as a solution to $\sum_{i=1}^{k} \alpha_i s_i = 0$, is said to be trivial if $(\alpha_1, \alpha_2, ..., \alpha_k) = (0, 0 ..., 0)$; otherwise it is nontrivial.*

Example 2.2. $V = C(\mathbb{R})$. Then $S_1 = \{\mathrm{sinx}, \mathrm{cosx}\}$ is linearly independent. $S_2 = \{\sin^2 x, \cos^2 x, 1\}$ is linearly dependent since

$$1 \cdot \sin^2 x + 1 \cdot \cos^2 x + (-1) \cdot 1 = 0 \quad (\forall x \in \mathbb{R})$$

Example 2.3. $V = \mathbb{R}[x]$. $S = \{1, x, x^2, x^3, ...\}$ is linearly independent. Pick x^{k_1}, $x^{k_2}, ..., x^{k_n} \in S, k_1 \leqslant \cdots \leqslant k_n$. To solve

$$\alpha_1 x^{k_1} + \cdots + \alpha_n x^{k_n} = 0$$

for all α_i, set $p(x) = \alpha_1 x^{k_1} + \cdots + \alpha_n x^{k_n}$, which is analytic on $\mathbb{R}$. Note that

$$\frac{d^{k_n}}{dx^{k_n}} p(x) = \alpha_n = 0$$

and thus $\alpha_1 x^{k_1} + \cdots + \alpha_{n-1} x^{k_{n-1}} = 0$. By differentiating $p(x)$ k_{n-1} times with respect to x, we have $\alpha_{n-1} = 0$. Repeat the differentiation n times and we may conclude that $\alpha_n = \alpha_{n-1} = \cdots = \alpha_1 = 0$. It follows that S is linearly independent.

Definition 2.4. *A subset S is basis of V if S pans V and S is linearly independent.*

Proposition 2.5. *Let $V = \operatorname{Span}(S)$ where $S = \{v_1, v_2 \ldots, v_m\} \subseteq V$. Then S contains a basis of V.*

Proof. If $S = \{v_1, v_2 \ldots, v_m\}$ is linearly independent then it is done. Suppose not, then $\sum_{i=1}^{m} \alpha_i v_i = 0$ has nontrivial solution.

By rearranging all v_i, assume that $\alpha_1 \neq 0$, then

$$v_1 = \left(-\frac{\alpha_2}{\alpha_1}\right) v_2 + \cdots + \left(-\frac{\alpha_m}{\alpha_1}\right) v_m$$

and thus $v_1 \in \operatorname{Span}\{v_2 \ldots, v_m\}$. By the proof of Lemma 1.11,

$$V = \operatorname{Span}\{v_1, v_2 \ldots, v_m\} = \operatorname{Span}\{v_2 \ldots, v_m\}$$

and let $S_1 = \{v_2 \ldots, v_m\}$. In this way, we can construct a sequence of subsets of S, i.e., $S_1, S_2, \ldots, S_k$, $k < m$, such that $S_1 \supseteq S_2 \supseteq \ldots \supseteq S_k$ and that S_k is linearly independent. Since

$$V = \operatorname{Span}(S) = \operatorname{Span}(S_1) = \operatorname{Span}(S_2) = \cdots = \operatorname{Span}(S_k)$$

$S_k \subseteq S$ is a basis of V.

□

Corollary 2.6. *If $V = \operatorname{Span}\{v_1 \ldots, v_m\}$ (finitely generated), then V has a basis.*

Lemma 2.7. *If $\{v_1, v_2 \ldots, v_n\}$ is a basis of V, then every $v \in V$ can be expressed uniquely as*

$$v = \sum_{i=1}^{n} \alpha_i v_i$$

Proof. Since $\{v_1, v_2 \ldots, v_n\}$ spans V, $v \in V$ can be written as

$$v = \sum_{i=1}^{n} \alpha_i v_i \quad (v_i \in \operatorname{Span}(S) = V)$$

Suppose

$$v = \sum_{i=1}^{n} \beta_i v_i$$

Rewrite

$$\sum_{i=1}^{n} \alpha_i v_i = \sum_{i=1}^{n} \beta_i v_i$$

as

$$\sum_{i=1}^{n} (\alpha_i - \beta_i)v_i = 0$$

Then it is clear that $(\alpha_i - \beta_i) = 0$ or $\alpha_i = \beta_i$ for all i since $\{v_1, v_2 ..., v_n\}$ is linearly independent.

□

The following theorem shows that every basis of a finitely generated vector space has the same size.

Theorem 2.8. *Let V be a finitely generated vector space, and $\{v_1, ..., v_m\}$, $\{w_1, ..., w_n\}$ are two bases of V. Then $m = n$. (This number n is called the dimension of V.)*

Proof. Suppose on contrary that $m \neq n$. Without loss of generality, assume $m < n$. Let $v_1 = \sum_{i=1}^{n} \alpha_i w_i$ such that α_i are not all zero. WLOG again assume $\alpha_1 \neq 0$. So by Lemma 1.11

$$v_1 \in \operatorname{Span}\{w_1, w_2 ..., w_n\} \backslash \operatorname{Span}\{w_2 ..., w_n\} \tag{2.1}$$

implies

$$w_1 \in \operatorname{Span}\{v_1, w_2 ..., w_n\} \backslash \operatorname{Span}\{w_2 ..., w_n\} \tag{2.2}$$

Now we claim that $\{v_1, w_2 ..., w_n\}$ is a basis of V and we shall prove this claim.

First we need to show that $\{v_1, w_2 ..., w_n\}$ spans V. Note that

$$\{w_1, ..., w_n\} \subseteq \operatorname{Span}\{v_1, w_2 ..., w_n\}$$

Then

$$V = \operatorname{Span}\{w_1, w_2 ..., w_n\} \subseteq \operatorname{Span}(\operatorname{Span}\{v_1, w_2 ..., w_n\}) = \operatorname{Span}\{v_1, w_2 ..., w_n\}$$

or $V \subseteq \operatorname{Span}\{v_1, w_2 ..., w_n\}$ by Lemma 1.11. Thus $\operatorname{Span}\{v_1, w_2 ..., w_n\} = V$. What left to be shown is the linear independence of this collection of vectors. Suppose for contradiction that $\beta_1 v_1 + \sum_{i=2}^{n} \beta_i w_i = 0$ has nontrivial solution. If $\beta_1 \neq 0$, we have

$$v_1 = \left(-\frac{\beta_2}{\beta_1}\right) w_2 + \cdots + \left(-\frac{\beta_n}{\beta_1}\right) w_n \in \operatorname{Span}\{w_2 ..., w_n\}$$

where a contradiction rises since $v_1 \in \operatorname{Span}\{w_1, w_2 ..., w_n\} \backslash \operatorname{Span}\{w_2 ..., w_n\}$. If $\beta_1 = 0$, then $\sum_{i=2}^{n} \beta_i w_i = 0$ has nontrivial solution, implying that $\{w_2 ..., w_n\}$ is linearly dependent. It follows that $\{w_1, ..., w_n\}$ is also linearly dependent, which is a contradiction. Thus the claim is proved.

Now we have $v_2 \in \operatorname{Span}\{v_1, w_2 ..., w_n\}$ and $v_2 = \gamma_1 v_1 + \gamma_2 w_2 + \cdots + \gamma_n w_n$ with some $\gamma_i \neq 0$. Note that $\gamma_i \neq 0$ for some $2 \leqslant i \leqslant n$, for otherwise $v_2 = \gamma_1 v_1$, $\gamma_1 \neq 0$ contradicts the linear independence of $\{v_1, ..., v_m\}$. By rearranging the order of $\gamma_2 w_2, ..., \gamma_n w_n$ we can assume that $\gamma_2 \neq 0$. Then we can express w_2 in terms of $v_1, v_2, w_3 ..., w_n$

$$w_2 = -\left(\frac{\gamma_1}{\gamma_2} v_1 + \frac{1}{\gamma_2} v_2 + \frac{\gamma_3}{\gamma_2} w_3 + \cdots + \frac{\gamma_n}{\gamma_2} w_n\right) \tag{2.3}$$

which leads to $w_2 \in \mathrm{Span}\{v_1, v_2, w_3, \ldots, w_n\}$. Similarly, $\{v_1, w_2 \ldots, w_n\} \subseteq \mathrm{Span}\{v_1, v_2, w_3, \ldots, w_n\}$, and thus $\{v_1, v_2, w_3, \ldots, w_n\}$ is a basis of V. Repeating this process we can conclude that $\{v_1, v_2, \ldots, v_m, w_{m+1}, \ldots, w_n\}$ is a basis of V. But any of $w_{m+1}, \ldots, w_n$ can be expressed in terms of $v_1, v_2, \ldots, v_m$, which implies that $\{v_1, v_2, \ldots, v_m, w_{m+1}, \ldots, w_n\}$ is linearly dependent. This contradiction establishes the theorem.

$\square$

Example 2.9. Let $V = \mathbb{F}^n$. Then $\{e_1, \ldots, e_n\}$ is a basis of V. In fact, the columns of any invertible $n \times n$ matrix forms a basis of V. That is, $\dim V = n$

Example 2.10. Let $V = M_{m \times n}(\mathbb{F})$. Then $\dim V = \mathrm{mn}$, and $\{E_{\mathrm{ij}} | \, 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n\}$ is a basis of V, where E_{ij} is $m \times n$ matrix with 1 at (i, j) entry and 0 at others.

Example 2.11. Let $V = \{A \in M_{n \times n}(\mathbb{R}) | \, A^T = A\}$, $W = \{B \in M_{n \times n}(\mathbb{R}) | \, B^T = -B\}$. Then $\dim V = {}^{n(n+1)}/_2$ and $\dim W = {}^{n(n-1)}/_2$. A rigorous proof may require at least two different bases to verify the dimension. Here we shed some light on how to understand it intuitively. Note that the space of all $n \times n$ matrices has dimension n^2, by the previous example. Because every symmetric matrix in V is determined by the entries on one side with respect to the diagonal and those on the diagonal, the number of entries, or E_{ij}, that determines the matrix, can be counted in the way

$$\frac{n^2 - n}{2} + n = \frac{n(n+1)}{2}$$

where ${}^{n^2-n}/_2$ stands for the number of entries on one side and n the number of those on the diagonal. Similarly, for W, it is plausible to conclude that the dimension is

$$\dim W = \dim V - n = \frac{n^2 - n}{2} = \frac{n(n-1)}{2}$$

Remark 2.12. Let $V = \mathbb{C}$. Then $V = \mathrm{Span}\{1, i\}$ is a vector space over $\mathbb{R}$, since all $z \in V$ is written as $z = a + bi$, $a, b \in \mathbb{R}$. Indeed, $\{1, i\}$ is a basis of V over $\mathbb{R}$, so $\dim V = 2$.

3 Operations on a Vector Space

Theorem 3.1. (Basis Extension) *Let V be a finite dimensional vector space and $\{v_1, \ldots, v_k\}$ a linearly independent set of vectors in V. Then this set of vectors can be extended to a basis $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ of V.*

Proof. Suppose $\dim V = n$ and $\{w_1, \ldots, w_n\}$ a basis of V. Consider $\{w_1, \ldots, w_n\} \cup \{v_1, \ldots, v_k\}$ which is linearly dependent. That is,

$$\sum_{i=1}^{n} \alpha_i w_i + \sum_{i=1}^{k} \beta_i v_i = 0$$

with some $\alpha_i \neq 0$. WLOG take $\alpha_1 \neq 0$. Then consider $\{w_2, \ldots, w_n\} \cup \{v_1, \ldots, v_k\}$. If it is linearly dependent, remove w_2. Repeat this process till a linearly independent set of vectors $S \cup \{v_1, \ldots, v_k\}$ such that $S \subseteq \{w_1, \ldots, w_n\}$. Rewrite $S = \{v_{k+1}, \ldots, v_n\}$, and we have $S' = \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ being a linearly independent set of vectors.

Now we claim that S' spans V. In fact, for all $w_i \in \{w_1, \ldots, w_n\}$, $w_i \in \mathrm{Span}(S')$. If not, consider the equation

$$\alpha w_i + \sum_{i=1}^{n} \beta_i v_i = 0 \tag{3.1}$$

Then $\alpha = 0$, for otherwise $w_i \in \mathrm{Span}(S')$ and $\beta_i = 0$ for all i, which implies that (3.1) has trivial solution merely. It follows that $\{w_1, \ldots, w_n\} \cup S \cup \{v_1, \ldots, v_k\}$ is linearly independent. This violates the maximality of S. Thus all $\{w_1, \ldots, w_n\} \subseteq \mathrm{Span}(S')$, and $\mathrm{Span}(S') = V$

□

Definition 3.2. *Let V be a vector space and W_1, W_2 subspaces of V.*

i. $W_1 \cap W_2 := \{w \in V \mid w \in W_1 \text{ and } w \in W_2\}$.

ii. $W_1 + W_2 := \{w_1 + w_2 \mid w_i \in W_i\}$

iii. If furthermore $W_1 \cap W_2 = \{0\}$, then $W_1 + W_2 = W_1 \oplus W_2$, which is called the direct product.

Proposition 3.3.

i. $W_1 \cap W_2$, $W_1 + W_2$ are vector spaces of V.

ii. $W_1 + W_2 = W_1 \oplus W_2$ if and only if any $w \in W_1 + W_2$ has a unique expression

$$w = w_1 + w_2$$

for some $w_1 \in W_1$, $w_2 \in W_2$.

Remark 3.4. We can define addition among a finite set of vector spaces $\{W_1, W_2, \ldots, W_k\}$ by $W_1 + W_2 + \cdots + W_k$. If $w_1 + w_2 + \cdots + w_k = 0$ implies $w_1 = \ldots = w_k = 0$, $W_1 + W_2 + \cdots + W_k = W_1 \oplus W_2 \oplus \cdots \oplus W_k$.

Proposition 3.5. (Complementation) *Let $W \subseteq V$ be a vector subspace of a finite-dimensional vector space V. Then there exists another subspace $W' \subseteq V$ such that*

$$W \oplus W' = V$$

Proof. Let $\dim W := k$ and $\dim V := n$. It is clear that $k \leqslant n$. Let $\{v_1, \ldots, v_k\}$ be a basis of W and by Theorem 3.1 we can extend it to a basis of V, denoted by $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$. Intuitively, take $W' := \mathrm{Span}\{v_{k+1}, \ldots, v_n\}$ and it remains to show that $W + W' = W \oplus W'$.

First we shall show that $W + W' = V$. It is obvious that $W + W' \subseteq V$. On the other hand, for any $v \in V$,

$$v = (\alpha_1 v_1 + \cdots + \alpha_k v_k) + (\alpha_{k+1} v_{k+1} + \cdots + \alpha_n v_n)$$

where $(\alpha_1 v_1 + \cdots + \alpha_k v_k) \in W$ and $(\alpha_{k+1} v_{k+1} + \cdots + \alpha_n v_n) \in W'$, implying $V \subseteq W + W'$. Thus $V = W + W'$.

To show that $W \cap W' = \{0\}$, let $v \in W \cap W'$. Since v can be uniquely expressed by a combination of $\{v_1, \ldots, v_k\}$ or $\{v_{k+1}, \ldots, v_n\}$, suppose $v = \sum_{i=1}^{k} \beta_i v_i = \sum_{i=k+1}^{n} \beta_i v_i$, and we have

$$\begin{aligned} v &= \beta_1 v_1 + \cdots + \beta_k v_k + 0 v_{k+1} + \cdots + 0 v_n \\ &= 0 v_1 + \cdots + 0 v_k + \beta_{k+1} v_{k+1} + \cdots + \beta_n v_n \\ &\in W \cap W' \end{aligned}$$

It follows by Lemma 2.7 that $\beta_i = 0$ for $i = 1, 2, \ldots, n$. Hence $v = 0$.

Here the proof completes.

□

4 Linear Transformation

Definition 4.1. *Let V, W be vector spaces. A mapping $T: V \longrightarrow W$ is a linear transformation if*

$$T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2)$$

for $\forall v_1, v_2 \in V$, $\forall \alpha, \beta \in \mathbb{R}$.

Example 4.2.

i. Let $A \in M_{m \times n}(\mathbb{R})$ (or $A \in \mathbb{R}^{m \times n}$). The mapping $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ defined as $T(x) = Ax$ for all $x \in \mathbb{R}^n$ is a linear transformation. Here we say $T \in L(\mathbb{R}^m, \mathbb{R}^n)$. If $m = n$, we denote it by $T \in L(\mathbb{R}^n)$.

ii. The mapping $T_1: \mathbb{R}[x] \longrightarrow \mathbb{R}[x]$ and $T_2: \mathbb{R}[x] \longrightarrow \mathbb{R}[x]$ defined as

$$\begin{aligned} T_1(p(x)) &= \frac{d}{dx} p(x) \\ T_2(p(x)) &= \int_0^x p(t) dt \end{aligned}$$

are both linear transformations.

iii. The mapping $T_1: M_{n \times n}(\mathbb{R}) \longrightarrow \mathbb{R}$ defined as

$$T_1(A) = \mathrm{trace}(A) := \sum_{i=1}^{n} a_{ii}$$

for all $A \in M_{n \times n}(\mathbb{R})$ is a linear transformation, while the mapping T_2: $M_{n \times n}(\mathbb{R}) \longrightarrow \mathbb{R}$ defined as

$$T_2(A) = \det[A]$$

is not a linear transformation.

Definition 4.3. *Let $T: V \longrightarrow W$ be a linear transformation.*

i. The kernel of T is defined as $\ker(T) = T^{-1}(0) := \{v \in V \,|\, T(v) = 0\}$.

ii. The image or range of T is defined as $T(V) := \{T(v) \in W \,|\, v \in V\}$.

Example 4.4.

i. Let $T \in L(\mathbb{R}^n)$ be a linear transformation with $T(x) = Ax$. Then

$$\ker(T) = \mathrm{Null}(A) := \{x \in \mathbb{R}^n \,|\, Ax = 0\}$$

where $\mathrm{Null}(A)$ is called the null space of A. The image of T is

$$T(V) = \mathrm{Col}(A) := \{Ax \,|\, x \in \mathbb{R}^n\} = \mathrm{Span}\{\text{the column vectors of } A\}$$

where $\mathrm{Col}(A)$ is called the column Space of A.

ii. Let $T: \mathbb{R}[x] \longrightarrow \mathbb{R}[x]$ be a linear transformation such that $T(p(x)) = \frac{d}{dx} p(x)$ for all $p(x) \in \mathbb{R}[x]$. Then

$$\ker(T) = \{p(x) = c \in \mathbb{R}[x] \,|\, c \in \mathbb{R}\}$$

and

$$T(\mathbb{R}[x]) = \mathbb{R}[x]$$

Proposition 4.5. *Let $T: V \longrightarrow W$ be a linear transformation. The kernel and image are themselves vector subspaces such that* $\ker \subseteq V$, $T(V) \subseteq W$.

Definition 4.6. *Let V, W be finite-dimensional vector spaces and $T: V \longrightarrow W$ a linear transformation. Then the rank of T is defined as*

$$\mathrm{rank}(T) := \dim T(V)$$

and the nullity of T as

$$\mathrm{nullity}(T) := \dim \ker(T)$$

Remark 4.7. The collection of all linear transformations between two vector spaces V and W (over field $\mathbb{F}$) is itself a vector space. To make it specific, we can define the vector addition and scalar multiplication as follows:

i. For $T, S \in L(V, W)$, the addition is defined as

$$(T + S)(v) = T(v) + S(v) \quad (v \in V)$$

ii. For any $T \in L(V, W)$, the scalar multiplication is defined as

$$(\gamma T)(v) = \gamma T(v) \quad (\gamma \in \mathbb{F}, v \in V)$$

In particular, if $V = \mathbb{R}^n$, $W = \mathbb{R}^m$, $L(V, W) = M_{m \times n}(\mathbb{R})$.

Proposition 4.8. *If* $\dim V = n$, $\dim W = m$, *then* $\dim L(V, W) = mn$.

The injectivity and surjectivity properties of a linear transformation T is characterised in the following proposition, using the kernel and image of it.

Proposition 4.9. *Let Let* $T: V \longrightarrow W$ *be a linear transformation between two finite-dimensional vector spaces over the field* $\mathbb{F}$.

i. T *is injective if and only if* $\ker(T) = \{0\}$, *i.e.,* $\operatorname{nullity}(T) = 0$.

ii. T *is surjective if and only if* $T(V) = W$, *i.e.,* $\operatorname{rank}(T) = \dim W$.

iii. *If* T *is bijective,* $T^{-1}: W \longrightarrow V$ *is also a linear transformation.*

Proof.

i. First suppose that $\ker(T) = \{0\}$. Let $v_1, v_2 \in V$ such that $T(v_1) = T(v_2)$. Then

$$0 = T(v_1) + (-T(v_2)) = T(v_1) + T(-v_2) = T(v_1 + (-v_2))$$

It follows that $v_1 + (-v_2) = 0$, i.e., $v_1 = v_2$. Thus T is injective.

Conversely, suppose that T is injective and $v \in \ker(T)$. Since

$$T(0) = T(v + (-v)) = T(v) + (-T(v)) = 0$$

we can conclude that $0 \in \ker(T)$. By injectivity $v = 0$. This completes the proof.

ii. It follows directly from the definition of surjective mapping.

iii. Note that $T(T^{-1}(w)) = w$ for all $w \in W$. It suffices to show that for any $\alpha, \beta \in \mathbb{F}$, $w_1, w_2 \in W$

$$T^{-1}(\alpha w_1 + \beta w_2) = T^{-1}(\alpha w_1) + T^{-1}(\beta w_2)$$

and

$$T^{-1}(\alpha w_1) = \alpha T^{-1}(w_1)$$

First, let $\alpha, \beta \in \mathbb{F}$, $w_1, w_2 \in W$ and $\alpha w_1 + \beta w_2 \in W$. Note that

$$T(T^{-1}(\alpha w_1) + T^{-1}(\beta w_2)) = T(T^{-1}(\alpha w_1)) + T(T^{-1}(\beta w_2)) = \alpha w_1 + \beta w_2$$

and

$$T(T^{-1}(\alpha w_1 + \beta w_2)) = \alpha w_1 + \beta w_2$$

Since T is a bijection, $T^{-1}(\alpha w_1 + \beta w_2) = T^{-1}(\alpha w_1) + T^{-1}(\beta w_2)$. It remains to show that $T^{-1}(\alpha w_1) = \alpha T^{-1}(w_1)$. In fact,

$$T(\alpha T^{-1}(w_1)) = \alpha T(T^{-1}(w_1)) = \alpha w_1 = T(T^{-1}(\alpha w_1))$$

and by the bijectivity of T we have $T^{-1}(\alpha w_1) = \alpha T^{-1}(w_1)$. Hence T^{-1} is a linear transformation.

□

Definition 4.10. (Isomorphism) *Let V, W be two vector spaces and $T\colon V \longrightarrow W$ an injective linear transformation. Then T is called an isomorphism from V to $T(V)$ and we say that T is isomorphic to $T(V)$, denoted by $V \cong T(V)$.*

Remark 4.11. It follows directly from the above definition that $V \cong W$ if T is bijective.

On the one hand, if $\dim V = \dim W = n < \infty$, then $V \cong W$. Let $\{v_1, \ldots, v_n\}$, $\{w_1, \ldots, w_n\}$ be bases of V and W respectively. It is plausible to construct T: $V \longrightarrow W$ such that $T(v_i) = w_i$ and $T(\alpha_i v_i) = \alpha_i T(v_i)$ for $i = 1, \ldots, n$. Then it follows that

$$T(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha_1 w_1 + \cdots + \alpha_n v_n \quad (\alpha_i \in \mathbb{F})$$

which is obviously a linear transformation.

Even though $V \cong W$, it is not necessary that every linear transformation T: $V \longrightarrow W$ is an isomorphism between V and W.

On the other hand, the size of basis is preserved under isomorphism; that is, $\dim V = \dim W$ if $V \cong W$ and they are both finite dimensional. It is because that every basis in V can be mapped to a basis in W through a bijective linear transformation. We shall show this in the following lemma before we move forward to the Rank-Nullity Theorem.

Lemma 4.12. *Let V, W be two finite-dimensional vector spaces. Suppose that $T\colon V \longrightarrow W$ is a isomorphism. Then $\dim V = \dim W$.*

Proof. Let $\dim V = n < \infty$ and let $\{v_1, \ldots, v_n\}$ be a basis of V. It suffices to show that $\{T(v_1), \ldots, T(v_n)\}$ is a basis of W.

Suppose that $\sum_{i=1}^n \alpha_i T(v_i) = 0$ for some $\alpha_i \in \mathbb{F}$. Then

$$\sum_{i=1}^n \alpha_i T(v_i) = \sum_{i=1}^n T(\alpha_i v_i) = 0$$

and thus $\sum_{i=1}^n \alpha_i v_i = 0, i = 1, \ldots, n$ by the injectivity of T (Proposition 4.9). Since $\{v_1, \ldots, v_n\}$ is linearly independent, $\alpha_1 = \cdots = \alpha_n = 0$. The linear independence of $\{T(v_1), \ldots, T(v_n)\}$ follows.

To show that $W = \operatorname{Span}\{T(v_1), \ldots, T(v_n)\}$, let $w \in W$ such that $w = T(v)$ for some $v \in V$. Since v can be uniquely expressed in terms of $\{v_1, \ldots, v_n\}$, assume that

$$v = \sum_{i=1}^n \beta_i v_i \quad (\beta_i \in \mathbb{F})$$

and

$$w = T(v) = T\left(\sum_{i=1}^n \beta_i v_i\right) = \sum_{i=1}^n \beta_i T(v_i)$$

Hence $\{T(v_1),...,T(v_n)\}$ spans W. The proof completes.

□

Theorem 4.13. (Rank-Nullity Theorem) *Let V, W be two vector spaces over the field $\mathbb{F}$ and $T\colon V\longrightarrow W$ a linear transformation. If $\dim V<\infty$, then*

$$\operatorname{rank}(T)+\operatorname{nullity}(T)=\dim V$$

Proof. If $T(V)=\{0\}$ or $T(V)=\emptyset$, the proof is trivial. Assume that $T(V)$ is not empty and $T(V)\neq\{0\}$. Then $V\neq\ker(T)$ and $\dim\ker(T)<\dim V$. By Proposition 3.5, there exists a subspace $V_1\subseteq V$ such that $V=V_1\oplus\ker(T)$, which leads to

$$\dim V=\dim V_1+\dim\ker(T)=\dim V_1+\operatorname{nullity}(T)$$

Now it suffices to prove that $\dim V_1=\dim T(V)=\operatorname{rank}(T)$.

Consider the linear transformation $T_1\colon V_1\longrightarrow W$ which satisfies

$$T_1(v)=T(v)\quad(\forall v\in V_1)$$

and $T(0)=0$. For every $v\in V_1$ with $T_1(v)=0=T(v)$, we have $v\in\ker(T)$. The injectivity follows immediately since $v\in\ker(T)\cap V_1=\{0\}$. Then T_1 is an isomorphism from V_1 to $T(V_1)$. By the preceding lemma, $\dim V_1=\dim T(V_1)$. It remains to show that $T(V_1)=T(V)$.

To show this, let $w\in T(V)$. If $w=0$, $w\in T(V_1)$. Otherwise, $w=T(v)$ for some $v\notin\ker(T)$ and $v\neq 0$, which implies $v\in V_1$ and $w\in T(V_1)$. Thus $T(V)\subseteq T(V_1)$. Since $T(V_1)\subseteq T(V)$ is obvious, $T(V_1)=T(V)$. Therefore, $\dim T(V_1)=\dim T(V)$ and this completes the proof.

□

The proof of Lemma 4.12 can be broken down into several results, which are presented in the following proposition.

Proposition 4.14. *If $T\colon V\longrightarrow W$ is an isomorphism, then*

- *i. the set $\{v_1,...,v_k\}$ is linearly independent in V if and only if $\{T(v_1),...,T(v_k)\}$ is linearly independent in W,*
- *ii. the set $\{v_1,...,v_k\}$ spans V if and only if $\{T(v_1),...,T(v_k)\}$ spans W,*
- *iii. $\{v_1,...,v_n\}$ forms a basis of V (provided that $\dim V=n$) if and only if $\{T(v_1),...,T(v_n)\}$ forms a basis of W, and*
- *iv. two finite-dimensional vector spaces are isomorphic if and only if they have the same dimension.*

A linear transformation $T\colon X\longrightarrow X$ is also called a linear operator on X. An important fact about the injectivity and surjectivity is that each of them implies the other, provided that T is a linear operator on some vector space X.

5 Change of Basis and Matrix Representation

Definition 5.1. (Coordinate Vector) *Let V be a n-dimensional vector space and $B=\{v_1,\ldots,v_n\}$ an ordered basis of V. The unique expression of every $v\in V$*

$$v=\alpha_1 v_1+\cdots+\alpha_n v_n \quad (\alpha_i\in\mathbb{F})$$

induces a mapping $[\cdot]_B\colon V\longrightarrow\mathbb{F}^n$ which maps each v to

$$[v]_B=(\alpha_1,\alpha_2,\ldots\alpha_n)\in\mathbb{F}^n$$

We call $[v]_B$ the coordinate vector, or simply, the coordinate of v.

Note that the mapping $[\cdot]_B\colon V\longrightarrow\mathbb{F}^n$ is well-defined for any given basis. In fact, if

$$[v]_B=(\alpha_1,\alpha_2,\ldots\alpha_n)=(\beta_1,\beta_2,\ldots,\beta_n)$$

then $\alpha_1 v_1+\cdots+\alpha_n v_n=\beta_1 v_1+\cdots+\beta_n v_n$ since the basis is ordered. This implies $\alpha_i=\beta_i$, $i=1,2,\ldots n$, by the uniqueness of the expression of v in terms of $\{v_1,\ldots, v_n\}$.

Example 5.2. Given $V=M_{2\times 2}(\mathbb{R})$ and an ordered basis

$$B=\left\{\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix},\begin{pmatrix}0&0\\1&0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix}\right\}$$

Then

$$\left[\begin{pmatrix}1&4\\2&3\end{pmatrix}\right]_B=(1,4,2,3)$$

However, the coordinate may change if a different basis is adopted. Let

$$B_1=\left\{\begin{pmatrix}0&1\\0&0\end{pmatrix},\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}0&0\\1&0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix}\right\}$$

then

$$\left[\begin{pmatrix}1&4\\2&3\end{pmatrix}\right]_B=(4,1,2,3)$$

Theorem 5.3. *Given a basis $B=\{v_1,\ldots,v_n\}$, the mapping $[\cdot]_B\colon V\longrightarrow\mathbb{F}^n$ is an isomorphism.*

Proof. We shall show that $[\cdot]_B\colon V\longrightarrow\mathbb{F}^n$ is a bijective linear transformation.

i. Let $v=\alpha_1 v_1+\cdots+\alpha_n v_n$ and $w=\beta_1 v_1+\cdots+\beta_n v_n$ be two vectors in V, and let $k\in\mathbb{F}$. Then

$$[kv]_B=(\alpha_1 k,\ldots,\alpha_n k)=k(\alpha_1,\ldots,\alpha_n)=k[v]_B$$

and

$$[v+w]_B=(\alpha_1+\beta_1,\ldots,\alpha_n+\beta_n)=(\alpha_1,\ldots,\alpha_n)+(\beta_1,\ldots,\beta_n)$$

It is clear that $[\cdot]_B$ is a linear transformation.

ii. Notice that $(\alpha_1, \ldots, \alpha_n) = (\beta_1, \ldots, \beta_n)$ implies $\alpha_i = \beta_i$, $i = 1, 2, \ldots n$. It leads to the same expression in terms of $\{v_1, \ldots, v_n\}$, i.e., the same coordinate vector in $\mathbb{F}^n$ is mapped from the same vector in V. Then the injectivity follows.

iii. For each vector $(\alpha_1, \ldots, \alpha_n)$ in $\mathbb{F}^n$, $\alpha_1 v_1 + \cdots + \alpha_n v_n$ is a linear combination of the basis B and thus $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ is a vector in V, which satisfies $[v]_B = (\alpha_1, \ldots, \alpha_n)$. The surjectivity follows.

□

Example 5.4. Let $V = P_3[x]$ with a basis $B = \{1, x, x^2, x^3\}$. By Proposition 4.14 and Theorem 5.3 to check if the set of vectors $\{1 + x^2, 3 - x^3, x - x^3\}$ is linearly independent, it suffices to check if the set of corresponding coordinates

$$\{(1, 0, 1, 0), (3, 0, 0, -1), (0, 1, 0, -1)\}$$

is linearly independent.

Since every basis is related to an isomorphism $[\cdot]_B$, it is of our interest to find the connection between the coordinates of a given vector v with respect to two distinct bases B_1 and B_2. In order to arrive at the proposition in which the change of basis is grounded, let us consider an example.

Example 5.5. Let $V = \mathbb{R}^n$ with a basis $B_1 = \{e_1, \ldots, e_n\}$, and let $v = (\alpha_1, \ldots, \alpha_n)$ be a vector in V, i.e.,

$$v = \alpha_1 e_1 + \cdots + \alpha_n e_n$$

and the coordinate of v with respect to B_1 is

$$[v]_{B_1} = (\alpha_1, \ldots, \alpha_n)$$

Now we change the basis B_1 to

$$B_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \right\}$$

which gives a different coordinate

$$[v]_{B_2} = (\alpha_1 - \alpha_2, \alpha_2 - \alpha_3, \ldots, \alpha_{n-1} - \alpha_n, \alpha_n)$$

Proposition 5.6. (Change of Basis) *Let $A = \{v_1, \ldots, v_n\}$ and $B = \{w_1, \ldots w_n\}$ be two ordered bases of a finite-dimensional vector space V. Define the change-of-basis matrix $\mathcal{C}_{B,A} := [\alpha_{ij}]$, the entries of which satisfy*

$$v_j = \sum_{i=1}^{n} \alpha_{ij} w_i \qquad (j = 1, 2, \ldots, n)$$

Then for any $v \in V$, the change of basis amounts to left-multiplying the matrix $\mathcal{C}_{B,A}$, i.e.,

$$[v]_B = \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} = \mathcal{C}_{B,A}\,[v]_A$$

where $v = h_1 v_1 + \cdots + h_n v_n = k_1 w_1 + \cdots + k_n w_n$. Furthermore, if we define

$$\mathcal{C}_{A,B} := [\beta_{rs}],$$

where

$$w_s = \sum_{r=1}^{n} \beta_{rs} v_r \qquad (s = 1, 2, \ldots, n)$$

then

$$[v]_A = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} = \begin{pmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1} & \beta_{n2} & \cdots & \beta_{nn} \end{pmatrix} \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} = \mathcal{C}_{A,B}\,[v]_B$$

i.e., $\mathcal{C}_{B,A} = (\mathcal{C}_{A,B})^{-1}$

Proof. To show $[v]_B = \mathcal{C}_{B,A}\,[v]_A$, it suffices to show $[v_j]_B = \mathcal{C}_{B,A}\,[v_j]_A$ for all j and the result then follows immediately from the linearity of the isomorphisms $[\cdot]_A$ and $[\cdot]_B$. So, for $v_j \in A$, $j = 1, 2, \ldots n$, we have $[v_j]_A = e_j$ and

$$\mathcal{C}_{B,A}\,[v_j]_A = \mathcal{C}_{B,A}\,e_j = (\alpha_{1j}, \alpha_{2j}, \ldots, \alpha_{nj}) = \left[\sum_{i=1}^{n} \alpha_{ij} w_i\right]_B = [v_j]_B$$

Then for any $v \in V$,

$$\begin{aligned} \mathcal{C}_{B,A}\,[v]_A &= \mathcal{C}_{B,A}\,[h_1 v_1 + \cdots + h_n v_n]_A \\ &= \mathcal{C}_{B,A} \sum_{j=1}^{n} h_j [v_j]_A \\ &= \sum_{j=1}^{n} h_j\, \mathcal{C}_{B,A}\,[v_j]_A \\ &= \sum_{j=1}^{n} h_j [v_j]_B \\ &= \left[\sum_{j=1}^{n} h_j v_j\right]_B \\ &= [v]_B \end{aligned}$$

Similarly, we can obtain that $[v]_A = \mathcal{C}_{A,B}[v]_B$. We shall show that $\mathcal{C}_{B,A} = (\mathcal{C}_{A,B})^{-1}$ to finish the proof. Note that for $j = 1, 2, \ldots, n$,

$$v_j = \sum_{i=1}^n \alpha_{ij} w_i = \sum_{i=1}^n \alpha_{ij} \left(\sum_{r=1}^n \beta_{ri} v_i \right) = \sum_{r=1}^n \left(\sum_{i=1}^n \alpha_{ij} \beta_{ri} \right) v_r$$

where

$$\sum_{i=1}^n \alpha_{ij} \beta_{ri} = \begin{cases} 1, r = j \\ 0, r \neq j \end{cases}$$

By multiplying the matrices $\mathcal{C}_{A,B}$ and $\mathcal{C}_{B,A}$, we may find that the (r, j)-th entry of $\mathcal{C}_{A,B}\mathcal{C}_{B,A}$ has the form

$$\sum_{i=1}^n \alpha_{ij} \beta_{ri}$$

This implies that the entries are all 1 on the diagonal and 0 elsewhere, i.e., $\mathcal{C}_{A,B}\mathcal{C}_{B,A} = I_n$, where I_n is the identity matrix. Analogous to the above, multiplying $\mathcal{C}_{B,A}$ and $\mathcal{C}_{A,B}$ yields the identity matrix, which completes the proof.

$\square$

Example 5.7. Let V, B_1, B_2 be the same as those in Example 5.5. Write $B_2 = \{w_1, \ldots, w_n\}$. Since $w_j = e_1 + \ldots + e_j$, the change-of-basis matrix from B_2 to B_1 is given by

$$\mathcal{C}_{B_1,B_2} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Then for $v \in V$ with $[v]_{B_2} = (\alpha_1 - \alpha_2, \alpha_2 - \alpha_3, \ldots, \alpha_{n-1} - \alpha_n, \alpha_n)$, its coordinate with respect to B_1 is given by

$$[v]_{B_1} = \mathcal{C}_{B_1,B_2}[v]_{B_2} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 - \alpha_2 \\ \alpha_2 - \alpha_3 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Definition 5.8. (Matrix Representation) *Let $T: V \longrightarrow W$ be a linear transformation, and let*

$$A = \{v_1, \ldots, v_n\}, B = \{w_1, \ldots, w_m\}$$

be bases of V and W respectively. The matrix representation of T with respect to A and B is defined as the matrix $T_{B,A} = [\alpha_{ij}] \in M_{m \times n}(\mathbb{F})$ such that

$$T(v_j) = \sum_{i=1}^m \alpha_{ij} w_i \qquad (j = 1, 2, \ldots, n)$$

Example 5.9. Let $V = P_3[x]$, $A = \{1, x, x^2, x^3\}$, and let V_A denote the vector space V equipped with the coordinate with respect to the basis A. Define T: $V_A \longrightarrow V_A$ as $T(p(x)) = p'(x)$ for all $p(x) \in V_A$. Note that

$$\begin{aligned} T(1) &= 0\cdot 1 + 0\cdot x + 0\cdot x^2 + 0\cdot x^3 \\ T(x) &= 1\cdot 1 + 0\cdot x + 0\cdot x^2 + 0\cdot x^3 \\ T(x^2) &= 0\cdot 1 + 2\cdot x + 0\cdot x^2 + 0\cdot x^3 \\ T(x^3) &= 0\cdot 1 + 0\cdot x + 3\cdot x^2 + 0\cdot x^3 \end{aligned}$$

Then we can find for T the matrix representation with respect to A (and A)

$$T_{A,A} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

If we choose another basis $A' = \{x^3, x^2, x, 1\}$, the matrix representation for T: $V_A \longrightarrow V_{A'}$ is given by

$$T_{A',A} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Here our observation is that the coordinate vectors before and after linear transformation admit a matrix multiplication. For instance, consider $T(2x^2 + 4x^3) = 4x + 12x^2$, where the corresponding coordinates are $(0, 0, 2, 4)$ and $(0, 4, 12, 0)$ with respect to A. Then

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 12 \\ 0 \end{pmatrix}$$

From Example 5.9, it is plausible to argue that linear transformation can be depicted as multiplying the coordinate vectors by the matrix representation. The next theorem will give a proof of this argument so that we can compute $T(v)$ by matrix multiplication.

Theorem 5.10. *Let $T: V \longrightarrow W$ be a linear transformation between two finite-dimensional vector spaces, and let A, B be the bases of V and W respectively. Then for any $v \in V$,*

$$T_{B,A}\,[v]_A = [T(v)]_B$$

which is illustrated in the proceeding diagram.

$$
\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, "{[\cdot]_A}"'] & W \arrow[d, "{[\cdot]_B}"] \\
\mathbb{F}^n \arrow[r, "\textit{left multiplying}", "T_{B,A}"'] & \mathbb{F}^m
\end{tikzcd}
$$

Figure 5.1. *Matrix representation where* $\dim V = n$, $\dim W = m$.

Proof. The proof is similar to that of Proposition 5.6 and is omitted here. □

Now it can be seen that the definition comes up as a generalisation of the change-of-basis method, i.e., Proposition 5.6 is a special case of Theorem 5.10. In fact, let $m = n$ in Definition 5.8, and consider, for any $v \in V$, the coordinate of v in terms of $\{v_1, \ldots, v_n\}$ and that of $T(v)$ in terms of $\{T(v_1), \ldots, T(v_n)\}$. In fact, the coordinate of v is preserved by the linearity of T, so, taking $\{T(v_1), \ldots, T(v_n)\}$ as a basis of W, T does not modify the coordinate. However, if we take another basis $\{w_1, \ldots, w_n\}$, we have to change the basis from $\{T(v_1), \ldots, T(v_n)\}$ to $\{w_1, \ldots, w_n\}$ within W. Then, by letting T be the identity mapping and $V = W$, $T_{B,A}$ is simply the change-of-basis matrix $\mathcal{C}_{B,A}$.

Proposition 5.11. (Functoriality) *Suppose that V, W, U are finite-dimensional vector spaces with the ordered bases A, B, C respectively. Let*

$$T: V \longrightarrow W, S: W \longrightarrow U$$

be two linear transformations. Then the composition of linear transformations correspond to multiplication of change-of-basis matrices, i.e.,

$$(S \circ T)_{C,A} = S_{C,B} T_{B,A}$$

Proof. Suppose that $\dim V = n$, $\dim W = m$, $\dim U = r$, and that

$$A = \{v_1, \ldots, v_n\}, B = \{w_1, \ldots, w_m\}, C = \{u_1, \ldots, u_r\}$$

To show $(S \circ T)_{C,A} = S_{C,B} T_{B,A}$, it suffices to show that each column of the matrix on the left hand side is equal to the corresponding one in the matrix on the right, i.e., $(S \circ T)_{C,A}\, e_j = S_{C,B} T_{B,A}\, e_j$, $j = 1, 2, \ldots, n$. Let the matrix representations be $T_{B,A} = [\tau_{ij}]_{m \times n}$, $S_{C,B} = [\varsigma_{ki}]_{r \times m}$, and $S_{C,B} T_{B,A} = [\gamma_{kj}]_{r \times n}$. Then

$$
\begin{aligned}
T(v_j) &= \sum_{i=1}^{m} \tau_{ij} w_i \\
S(w_i) &= \sum_{k=1}^{r} \varsigma_{ki} u_k
\end{aligned}
$$

Note that $e_j = [v_j]_A$. Then we shall show that $(S \circ T)_{C,A}\,[v_j]_A = S_{C,B}\,T_{B,A}\,[v_j]_A$, $j = 1, 2, ..., n$. For $v_j \in A$, $j = 1, 2, ..., n$,

$$\begin{aligned}
(S \circ T)_{C,A}[v_j]_A &= [(S \circ T)(v_j)]_C \\
&= [S(T(v_j))]_C \\
&= \left[S\left(\sum_{i=1}^{m} \tau_{ij} w_i\right)\right]_C \\
&= \left[\sum_{i=1}^{m} \tau_{ij}\, S(w_i)\right]_C \\
&= \left[\sum_{i=1}^{m} \tau_{ij}\left(\sum_{k=1}^{r} \varsigma_{ki} u_k\right)\right]_C \\
&= \left[\sum_{k=1}^{r} \left(\sum_{i=1}^{m} \tau_{ij}\varsigma_{ki}\right) u_k\right]_C \\
&= \sum_{k=1}^{r} \left(\sum_{i=1}^{m} \tau_{ij}\varsigma_{ki}\right) [u_k]_C \\
&= \sum_{k=1}^{r} \gamma_{kj} e_k
\end{aligned}$$

where $\sum_{k=1}^{r} \gamma_{kj} e_k$ is the j-th column of $S_{C,B}\,T_{B,A}$. The result follows immediately.

□

Consider two identity mappings S and T on V, and their matrix representations $S_{A',A}$ and $T_{A,A'}$ where A and A' are two bases of V. It can be easily seen that $S_{A',A}$ and $T_{A,A'}$ are also two change-of-basis matrices on V. Then, denoting them by $\mathcal{C}_{A',A}$ and $\mathcal{C}_{A,A'}$,

$$\mathcal{C}_{A',A}\,\mathcal{C}_{A,A'} = S_{A',A}\,T_{A,A'} = (S \circ T)_{A',A'}$$

by Proposition 5.11, where $(S \circ T)_{A',A'}$ is the matrix representation of an identity mapping with respect to the same basis A', i.e., the identity matrix. This admits the result $\mathcal{C}_{A',A}\,\mathcal{C}_{A,A'} = I$ in Proposition 5.6.

Proposition 5.12. *Let $T: V \longrightarrow W$ be a linear transformation and let* $\dim V = n$, $\dim W = m$. *If A and A' are two ordered bases of V, and B and B' are tow ordered bases of W, then the change-of-basis matrices admit the relation*

$$T_{B',A'} = \mathcal{C}_{B',B}\,T_{B,A}\,\mathcal{C}_{A,A'}$$

Proof. It can be done simply by considering the j-th column of the matrices on the left hand side and the right hand side and showing them equal to each other. □

Remark 5.13. If T is a linear operator on V, and if A, A' are again two ordered bases of V, then by the preceding proposition,

$$T_{A',A'} = \mathcal{C}_{A',A} T_{A,A} \mathcal{C}_{A,A'} = (\mathcal{C}_{A,A'})^{-1} T_{A,A'} \mathcal{C}_{A,A'}$$

which implies that $T_{A',A'}$ is similar to $T_{A,A}$.

6 Quotient Spaces

This section is aiming at dividing a big vector space in to some "slices" by choosing certain representatives to characterise these "slices". To do so, it is natural to induce a equivalence relation on this vector space and divide the space into equivalence classes. Here we adopt "congruence" as the equivalence relation.

Let V be a vector space and W a subspace of V. For any $\alpha, \beta \in V$, we say that α is **congruent to** β **modulo** W if $\alpha - \beta \in W$, denoted by

$$\alpha \equiv \beta \pmod{W}$$

It is easy to verify that congruence modulo W is an equivalence relation. So, we shall then study the equivalence classes induced by this equivalence relation.

Definition 6.1. *Let V be a vector space and W a subspace of V. For any $v \in V$, the (right) coset of v relative to W is defined as*

$$v + W := \{v + w \mid w \in W\}$$

Remark 6.2. A representative of a coset $v + W$ is a vector $v + w \in v + W$.

It is not quite clear how these "cosets" are related to the equivalence classes we would like to study. This is illustrated in the next lemma.

Lemma 6.3. *Let W be subspace of a vector space V and $v_1, v_2 \in V$. Then $v_1 + W = v_2 + W$ if and only if v_1 is congruent to v_2 modulo W, i.e., $v_1 - v_2 \in W$.*

Proof. First, suppose $v_1 + W = v_2 + W$. Then $v_1 + w_1 = v_2 + w_2$ for some w_1, $w_2 \in W$, and thus $v_1 - v_2 = w_2 - w_1 \in W$.

Conversely, suppose $v_1 - v_2 = w' \in W$. Then $v_1 + 0 = v_2 + w'$. For all $v_1 + w \in v_1 + W$, we have

$$v_1 + w = v_1 + 0 + w = v_2 + w' + w \in v_2 + W.$$

Thus $v_1 + W \subseteq v_2 + W$. Similarly, $v_2 + W \subseteq v_1 + W$. Thus $v_1 + W = v_2 + W$. □

Remark 6.4. $v + W = W$ if and only if $v \in W$.

Exercise 6.1. If $v_1 - v_2 \notin W$, then $(v_1 + W) \cap (v_2 + W) = \emptyset$.

Definition 6.5. *The quotient space V/W is the collection of all cosets, i.e.,*

$$V/W := \{v + W \mid v \in V\}$$

Remark 6.6. $0_{V/W} = 0 + W$.

Now we define addition on V/W by

$$(v_1 + W) + (v_2 + W) := (v_1 + v_2) + W$$

and scalar multiplication by

$$\alpha(v_1 + W) = \alpha v_1 + W$$

What follows is to check if they are well defined, for otherwise we are not able to say anything more about the structure relying on these operations. Here one comment should be made. To show that certain binary operations, say, the addition and multiplication, are well defined, one should verify that these operations yields exactly one "output" for each pair of "inputs". For instance, when checking the addition, we shall find an arbitrary pair of "inputs", say, $(v_1 + W)$ and $(v_2 + W)$, and compare the "output" – $(v_1 + v_2) + W$ – of them with that of another pair, say, $v_1' + W$ and $v_2' + W$. Note that the latter pair of "inputs" should be equivalent to the former up to the equivalence relation defined previously; that is, WLOG,

$$v_1 \equiv v_1', v_2 \equiv v_2' \pmod{W}$$

which means $v_1 + W = v_1' + W$, $v_2 + W = v_2' + W$ as well.

Proposition 6.7. *The addition and scalar multiplication are well-defined, and it gives a vector space structure on V/W.*

Proof. Suppose $v_1 + W = v_1' + W$, $v_2 + W = v_2' + W$. By Lemma 6.3,

$$v_1 - v_1' = w_1, v_2 - v_2' = w_2 \quad (w_1, w_2 \in W)$$

It leads to

$$(v_1 + v_2) - (v_1' + v_2') = w_1 + w_2$$

which implies

$$(v_1 + v_2) + W = (v_1' + v_2') + W$$

For scalar multiplication, if $v_1 + W = v_1' + W$, then $v_1 - v_1' = w_1 \in W$. For any scalar $\alpha \in \mathbb{F}$,

$$\alpha v_1 - \alpha v_1' = \alpha(v_1 - v_1') \in W$$

and thus $\alpha v_1 + W = \alpha v_1' + W$.

It is trivial to check each of the axioms of vector space, and therefore omitted here.

□

There is a natural mapping from V onto V/W; that is, the linear transformation $\pi_W: V \longrightarrow V/W$ with $\pi_W(v) = v + W$ for any $v \in V$. This canonical mapping is called **quotient mapping** (or **quotient transformation**). The kernel of π_W is precisely W.

Proposition 6.8. *The canonical mapping $\pi_W: V \longrightarrow V/W$ is surjective linear transformation with its kernel being precisely W.*

Proof. For any $v_0 + W \in V/W$, there exists $v_0 \in V$ such that $\pi_W(v_0) = v_0 + W$.

To show that it is a linear transformation, check if $\pi_W(\alpha v_1 + \beta v_2) = \alpha \pi_W(v_1) + \beta \pi_W(v_2)$.

Note that $0_{V/W} = W$. By Remark 6.4, $\pi_W(v) \in W$ if and only if $v \in W$. $\square$

Remark 6.9. Let $A: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear transformation. To solve $Ax = b$, it suffices to

i. find $\ker(A) = \mathrm{Null}(A)$ $(Ax = 0)$,

ii. find a particular solution $Ax_0 = b$, and

iii. solutions of $Ax = b$ form the collection $x_0 + \ker(A)$, which is a vector in $\mathbb{R}^n / \ker(A)$.

So it is of our interest to study the quotient space $\mathbb{R}^n / \ker(A)$.

Proposition 6.10. (Universal Property 1) *Suppose that $T: V \longrightarrow W$ is a linear transformation , and that V' is a subspace of* $\ker(T)$*. Then the mapping $\bar{T}$: $V/V' \longrightarrow W$ which maps $v + V'$ to $T(v)$ is a well-defined linear transformation.*

Proof. To show that $\bar{T}$ is well defined, suppose $v_1 + V' = v_2 + V'$. Then there exists $v' \in V'$ such that $v_1 - v_2 = v'$. It follows that

$$T(v_1) - T(v_2) = T(v_1 - v_2) = T(v') = 0$$

or $T(v_1) = T(v_2)$. Thus $\bar{T}(v_1 + V') = T(v_1) = T(v_2) = \bar{T}(v_2 + V')$.

Take $v_1, v_2 \in V$ and $\alpha, \beta \in \mathbb{F}$. Since $\bar{T}(v_1 + V') + \bar{T}(v_2 + V') = T(v_1) + T(v_2)$, and

$$\begin{aligned} \bar{T}((v_1 + V') + (v_2 + V')) &= \bar{T}((v_1 + v_2) + V') \\ &= T(v_1 + v_2) \\ &= T(v_1) + T(v_2) \end{aligned}$$

it is clear that $\bar{T}((v_1 + V') + (v_2 + V')) = \bar{T}(v_1 + V') + \bar{T}(v_2 + V')$. Then $\bar{T}$ is closed under addition. For scalar multiplication, we have

$$\begin{aligned} & \bar{T}(\alpha(v_1 + V')) \\ = \ & \bar{T}(\alpha v_1 + V') \\ = \ & T(\alpha v_1) \\ = \ & \alpha T(v_1) \\ = \ & \alpha \bar{T}(v_1 + V') \end{aligned}$$

Therefore, $\bar{T}$ is a linear transformation.

$\square$

Theorem 6.11. (First Isomorphism Theorem) *Let $T: V \longrightarrow W$ be surjective linear transformation. Then the mapping $\bar{T}: V/\ker(T) \longrightarrow W$ with $\bar{T}(v + \ker(T)) = T(v)$ for any $v \in V$ is an isomorphism.*

Proof. Let $v_1, v_2 \in V$ be two vectors satisfying

$$\bar{T}(v_1 + \ker(T)) = \bar{T}(v_2 + \ker(T))$$

The injectivity follows immediately since

$$T(v_1) = \bar{T}(v_1 + \ker(T)) = \bar{T}(v_2 + \ker(T)) = T(v_2)$$

implies $v_1 - v_2 \in \ker(T)$, i.e., $v_1 \equiv v_2 \pmod{\ker(T)}$.

Since T is surjective, then for any $w \in W$, there exists a vector $v \in V$ such that

$$w = T(V) = \bar{T}(v + \ker(T))$$

which implies $\bar{T}$ is surjective as well.

$\square$

Definition 6.12. (Universal Property For Quotients) *Let V be a vector space and V' a subspace of it. Consider the collection*

$$O_{V,V'} = \{T: V \longrightarrow W \,|\, T \in L(V, W) \text{ and } V' \leqslant \ker(T)\}$$

We say that $\phi \in O_{V,V'}$ with $\phi: V \longrightarrow U$ satisfies the universal property (or ϕ is a universal object) if for every $T \in O_{V,V'}$ mapping V to some W, there exists a unique linear transformation $\bar{T}: U \longrightarrow W$ such that $T = \bar{T} \circ \phi$, i.e., the following diagram commutes:

$$\begin{tikzcd}
V \arrow[rr, "T"] \arrow[dr, "\phi"'] & & W \\
& U \arrow[ur, "\bar{T}"'] &
\end{tikzcd}$$

Figure 6.1. $T = \bar{T} \circ \phi$

Theorem 6.13. (Universal Property 2)

i. The quotient mapping $\pi_{V'} \in O_{V,V'}$ is a universal object.

ii. If $\phi \in O_{V,V'}$ is a universal object that maps V to U, then U is isomorphic to V/V'.

Proof.

i. Let $T: V \longrightarrow W$ be such that $V' \leqslant \ker(T)$. Define $\bar{T}: V/V' \longrightarrow W$ as in Proposition 6.10. Then $\bar{T}$ is a well-defined linear transformation satisfying $T = \bar{T} \circ \phi$ by Universal Property 1. What follows is to show the uniqueness. Suppose there exists another mapping $\bar{S}: V/V' \longrightarrow W$ such that $T = \bar{S} \circ \phi$. Then, for any $v \in V$, $T(v) = \bar{S} \circ \pi_{V'}(v) = \bar{S}(v + V')$ since $\pi_{V'}$ is surjective. Therefore,
$$\bar{T}(v + V') = T(v) = \bar{S}(v + V')$$
for all $v \in V$.

ii. Suppose $\phi: V \longrightarrow U$ is a universal object. Note that $\pi_{V'} \in O_{V,V'}$ satisfies the universal property by the result above. Then there exists a unique linear transformation $\bar{\pi}_{V'}: U \longrightarrow V/V'$ such that $\phi = \bar{\pi}_{V'} \circ \pi_{V'}$, i,e, the following diagram commutes:

$$\begin{tikzcd} V \arrow[rr, "\pi_{V'}"] \arrow[dr, "\phi"'] & & V/V' \\ & U \arrow[ur, "\bar{\pi}_{V'}"'] & \end{tikzcd}$$

Figure 6.2. $\pi_{V'} = \bar{\pi}_{V'} \circ \phi$

Similarly, there exists a unique linear transformation $\bar{\phi}: V/V' \longrightarrow U$ such that $\pi_{V'} = \phi \circ \bar{\phi}$, which is shown in the diagram below:

$$\begin{tikzcd} V \arrow[rr, "\phi"] \arrow[dr, "\pi_{V'}"'] & & U \\ & V/V' \arrow[ur, "\bar{\phi}"'] & \end{tikzcd}$$

Figure 6.3. $\phi = \bar{\phi} \circ \pi_{V'}$

Combining $\phi = \pi_{V'} \circ \bar{\pi}_{V'}$ and $\pi_{V'} = \phi \circ \bar{\phi}$ together we have
$$\begin{aligned} \phi &= \bar{\phi} \circ \pi_{V'} = \bar{\phi} \circ \bar{\pi}_{V'} \circ \phi \\ \pi_{V'} &= \bar{\pi}_{V'} \circ \phi = \bar{\pi}_{V'} \circ \bar{\phi} \circ \pi_{V'} \end{aligned}$$

which implies that $\bar{\phi} \circ \bar{\pi}_{V'}$ and $\bar{\pi}_{V'} \circ \bar{\phi}$ are both identity mappings. WLOG, consider $\bar{\pi}_{V'} \circ \bar{\phi}$. Since $\pi_{V'}$ is surjective, $\bar{\pi}_{V'} \circ \bar{\phi}$ is surjective. This obviously leads to the surjectivity of $\bar{\pi}_{V'}$. As for $\bar{\phi}$, let $v_1, v_2 \in V$ satisfy $\bar{\phi}(v_1 + V') = \bar{\phi}(v_2 + V')$. It follows that

$$\begin{aligned} \bar{\pi}_{V'} \circ \bar{\phi}(v_1 + V') &= \bar{\pi}_{V'}(\bar{\phi}(v_1 + V')) \\ &= \bar{\pi}_{V'}(\bar{\phi}(v_2 + V')) \\ &= \bar{\pi}_{V'} \circ \bar{\phi}(v_2 + V') \end{aligned}$$

and thus $v_1 + V = v_2 + V'$ for $\bar{\pi}_{V'} \circ \bar{\phi}$ is an identity mapping. So, $\bar{\phi}$ is injective.

Analogous to the above, if we consider $\bar{\phi} \circ \bar{\pi}_{V'}$, we can show that $\bar{\phi}$ is surjective and that $\bar{\pi}_{V'}$ is injective. Now it can be concluded that $\bar{\phi}$ and $\bar{\pi}_{V'}$ are both bijective linear transformations, i.e., isomorphisms, between U and V/V'.

□

7 Dual Spaces

Definition 7.1. *Let V be a vector space over the field $\mathbb{F}$. A linear transformation $T: V \longrightarrow \mathbb{F}$ is called a linear functional on V. The collection of all linear functionals on V is called the dual vector space, or simply dual space V^*, i.e.,*

$$V^* := \mathrm{Hom}_{\mathbb{F}}(V, \mathbb{F}) = \{f: V \longrightarrow \mathbb{F} | f \in L(V, \mathbb{F})\}$$

Remark 7.2. V^* is a vector space over the field $\mathbb{F}$. It is left for the reader to check.

Example 7.3.

i. Let $V = \mathbb{R}^n$. The linear transformation $\phi_i: V \longrightarrow \mathbb{R}$ with

$$\phi_i(x) = x_i, i = 1, 2, ..., n$$

is a linear functional and thus in V^*.

ii. Let $V = \mathbb{F}[x]$. Then $\phi: V \longrightarrow \mathbb{F}$ with

$$\phi(p(x)) = p(1)$$

and $\varphi: V \longrightarrow \mathbb{F}$ with

$$\varphi(p(x)) = \int_0^1 p(x) \mathrm{d}x$$

are both linear functionals and thus in V^* as well.

iii. Let $V = M_{n \times n}(\mathbb{F})$. Then $\mathrm{tr}: V \longrightarrow \mathbb{F}$ with

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$$

for any $A=[a_{ij}]\in V$ is a linear functional and thus in V^*.

Definition 7.4. *Let V be a vector space with the basis $B=\{v_i|i\in I\}$ where I can be a finite, countably infinite, or uncountable index set. Define*

$$B^*:=\{f_i: V\longrightarrow \mathbb{F}|i\in I\}$$

by

$$f_i(v_j)=\delta_{ij}=\begin{cases}1, i=j\\ 0, i\neq j\end{cases}$$

where we denote $1_{\mathbb{F}}$ and $0_{\mathbb{F}}$ by 1 and 0 respectively for simplicity. Extend all f_i linearly, i.e., for any $\sum_{j=1}^n \alpha_j v_j\in V$,

$$f\left(\sum_{j=1}^n \alpha_j v_j\right)=\sum_{j=1}^n \alpha_j f_i(v_j)$$

It is clear that all f_i are well-defined linear functionals and thus all in V^*.

Lemma 7.5.

i. B^* is linearly independent.

ii. If V is finite-dimensional, then B^* is a basis of V^*.

Proof.

i. Take an arbitrary collection $\{f_1, ..., f_k\}$ and we shall show that it is linearly independent. Suppose

$$\sum_{r=1}^k \alpha_r f_r=0_{V^*}$$

where 0_{V^*} is a zero mapping which maps all $v\in V$ to $0\in\mathbb{R}$. Then

$$\sum_{r=1}^k \alpha_r f_r(v_1)=0$$

Since $f_r(v_1)=0$ when $r\neq 1$, we can conclude that $\alpha_1=0$. Similarly, one can show that

$$\alpha_2=\cdots=\alpha_k$$

and B^* is linearly independent.

ii. Suppose $B=\{v_1,...v_n\}$ and $B^*=\{f_1,...f_n\}$. For any $f\in V^*$, we shall show that

$$f(v_j)=\sum_{i=1}^n \alpha_i f_i(v_j)$$

for some $\alpha_i \in \mathbb{F}$, $i = 1, 2, ..., n$. Note that $f_i(v_j) = 0$ for $i \neq j$ and $f_i(v_j) = 1$ for $i = j$. It follows that $f(v_j) = \alpha_j$. Then it is plausible to set $\alpha_i = f(v_i)$ and construct

$$g(v) := \sum_{i=1}^{n} \alpha_i f_i(v) = \sum_{i=1}^{n} f(v_i) \cdot f_i(v) \in \operatorname{Span}(B^*)$$

where $f_i \in B^*$. Then

$$g(v_j) = \sum_{i=1}^{n} f(v_i) \cdot f_i(v_j) = f(v_j)$$

for $j = 1, 2, ..., n$, i.e., $g(v) = f(v)$ for all $v \in V$. Thus $f = g \in \operatorname{Span}(B^*)$ and B^* spans V.

□

Remark 7.6. From the proof we may see that if $f \in V^*$ and $B^* = \{f_1, ... f_n\}$ is a basis of V^*, f must have the form

$$f = \sum_{i=1}^{n} \alpha_i f_i = \sum_{i=1}^{n} f(v_i) f_i$$

Corollary 7.7. *If* $\dim V = n$, *then* $\dim V^* = n$ *and* $V \cong V^*$.

Remark 7.8. An isomorphism between V and V^* is the mapping from v_i to f_i and it depends on the choice of basis B in V.

Example 7.9. The second part of Lemma 7.5 does not hold for infinite-dimensional vector space V. Here is an counterexample. Let $V = \mathbb{F}[x]$. $B = \{1, x, x^2, ...\}$ is a basis of V. $B^* = \{\phi_0, \phi_1, \phi_2, ...\}$ with

$$\phi_i(x^j) = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}$$

is then a basis of V^*. Consider $\phi \in V^*$ with $\phi(p(x)) = p(1)$. Note that $\phi(x^n) = 1$ for all $n \in \mathbb{N}$. Then

$$\phi = \sum_{n=0}^{\infty} \phi(x^n) \cdot \phi_n = \sum_{n=0}^{\infty} \phi_n$$

which is an infinite sum and cannot be spanned by B^*.

8 Annihilators

Definition 8.1. *let* V *be a vector space and* $S \subseteq V$ *a subset. The annihilator of* S *is defined as*

$$\operatorname{Ann}(S) = \{f \in V^* \mid f(s) = 0_{\mathbb{F}}, \forall s \in S\}$$

Example 8.2. Let $V=\mathbb{R}^4$ with a basis $B=\{e_1,e_2,e_3,e_4\}$ the dual space V^* with a basis $B^*=\{f_1,f_2,f_3,f_4\}$. Take a subset $S=\{e_3,e_4\}\subseteq V$. Then $f_1\in\operatorname{Ann}(S)$ since $f_1(e_3)=0$ and $f_1(e_4)=0$. In fact,

$$\alpha f_1+\beta f_2\in\operatorname{Ann}(S)$$

for any $\alpha,\beta\in\mathbb{R}$ and $\operatorname{Ann}(S)=\operatorname{Span}\{f_1,f_2\}$.

Proposition 8.3. *Let V be a vector space with dual space V^*.*

i. $\operatorname{Ann}(S)$ *is a vector subspace of* V^**;*

ii. If $W_1,W_2\subseteq V$ *satisfy* $W_1\subseteq W_2$*, then* $\operatorname{Ann}(W_2)\subseteq\operatorname{Ann}(W_1)$*;*

iii. $\operatorname{Ann}(S)=\operatorname{Ann}(\operatorname{Span}(S))$*;*

iv. If V is finite-dimensional, and W is a subspace of V. Then $\dim W+\dim\operatorname{Ann}(W)=\dim V$.

Proof.

i. Suppose $f,g\in\operatorname{Ann}(S)$. Then

$$(\alpha f+\beta g)(s)=\alpha f(s)+\beta g(s)=0$$

and thus $\alpha f+\beta g\in\operatorname{Ann}(S)$.

ii. Take $f\in\operatorname{Ann}(W_2)$ and we have $f(w)=0$ for all $w\in W_2$, which in turn implies $f(w)=0$ for all $w\in W_1$ since $W_1\subseteq W_2$. This shows that $f\in\operatorname{Ann}(W_1)$ and $\operatorname{Ann}(W_2)\subseteq\operatorname{Ann}(W_1)$

iii. Since $S\subseteq\operatorname{Span}(S)$, by ii we have $\operatorname{Ann}(\operatorname{Span}(S))\subseteq\operatorname{Ann}(S)$. It suffices to show that

$$\operatorname{Ann}(S)\subseteq\operatorname{Ann}(\operatorname{Span}(S)) \tag{8.1}$$

For any $f\in\operatorname{Ann}(S)$ and any $\sum_{i=1}^n k_i s_i\in\operatorname{Span}(S)$, we have

$$f\left(\sum_{i=1}^n k_i s_i\right)=\sum_{i=1}^n k_i f(s_i)=0$$

Hence $f\in\operatorname{Ann}(\operatorname{Span}(S))$ and (8.1) follows.

iv. Suppose

$$\dim V=n,\quad \dim W=k\leqslant n$$

Let $\{v_1,\ldots,v_k\}$ be a basis of W. By basis extension, there exists $B=\{v_1,\ldots,v_k,v_{k+1},\ldots,v_n\}$ as a basis of V, and $B^*=\{f_1,\ldots,f_n\}$ as a basis of V^*. By Lemma 7.5

$$\dim\operatorname{Span}\{v_{k+1},\ldots,v_n\}=\dim\operatorname{Span}\{f_{k+1},\ldots,f_n\}.$$

Since

$$\dim W+\dim\operatorname{Span}\{v_{k+1},\ldots,v_n\}=\dim V \tag{8.2}$$

we claim that $\{f_{k+1}, \ldots, f_n\}$ is a basis of $\mathrm{Ann}(W)$ and it suffices to prove this claim.

It is clear that $\{f_{k+1}, \ldots, f_n\}$ is a linearly independent subset of $\mathrm{Ann}(W)$, because

$$f_{k+1}(v_j) = f_{k+2}(v_j) = \cdots = f_n(v_j) = 0$$

for $j = 1, \ldots, k$ (which means that $f_{k+1}, \ldots, f_n \in \mathrm{Ann}(W)$) and $\{f_1, \ldots, f_n\}$ is a basis of V^* (then $\{f_{k+1}, \ldots, f_n\}$ as a subset is linearly independent). What remains to be shown is that $\{f_{k+1}, \ldots, f_n\}$ spans $\mathrm{Ann}(W)$. For all $f \in V^*$, we have by Remark 7.6

$$f = \sum_{i=1}^{n} \alpha_i f_i = \sum_{i=1}^{n} f(v_i) f_i \tag{8.3}$$

If f is furthermore in $\mathrm{Ann}(W)$, $f(v_i) = 0$ for $i = 1, 2, \ldots, k$. Then

$$f = \sum_{i=1}^{k} f(v_i) f_i + \sum_{i=k+1}^{n} f(v_i) f_i = \sum_{i=k+1}^{n} f(v_i) f_i$$

and the claim is proved. Now the whole proof completes. □

For any subspace W of V, where V is finite-dimensional,

$$\dim \mathrm{Ann}(W) = \dim V - \dim W$$

and

$$\dim (V/W)^* = \dim V/W = \dim V - \dim W.$$

In fact, $(V/W)^* \cong \mathrm{Ann}(W)$. We shall show this in the next section, by constructing an isomorphism explicitly.

9 Adjoint Map

Definition 9.1. *Let $T \in L(V, W)$. The adjoint of T is defined by*

$$T^*: W^* \longrightarrow V^*$$

such that for any $f \in W^*$,

$$T^*(f) = f \circ T = fT.$$

Remark 9.2. It is easy to see that T^* itself is a linear transformation since it is the composition of two linear transformations. More explicitly, for any f, $g \in W^*$ and $a, b \in \mathbb{F}$,

$$[T^*(af + bg)](v) = [(af + bg) \circ T](v) = af(T(v)) + bg(T(v))$$

and by the definition of T^* it yields that

$$af(T(v)) + bg(T(v)) = aT^*(f)(v) + bT^*(g)(v) = [aT^*(f) + bT^*(g)](v)$$

for all $v \in V$. It is clear that $T^*(af + bg) = aT^*(f) + bT^*(g)$.

Proposition 9.3. *Let $T \in L(V, W)$.*

i. If T is injective, then T^* is surjective.

ii. If T is surjective, then T^* is injective.

Proof. The proof is omitted here. □

Proposition 9.4. *Let $T: V \longrightarrow W$ be a linear transformation and $A = \{v_1, \ldots, v_n\}$, $B = \{w_1, \ldots, w_m\}$ the bases of V and W respectively. If*

$$A^* = \{f_1, \ldots, f_n\}, B^* = \{g_1, \ldots, g_m\}$$

are the bases of the dual spaces V^* and W^* respectively, then the adjoint T^*: $W^* \longrightarrow V^*$ admits a matrix representation

$$T^*_{A^*, B^*} = (T_{B,A})^{\mathrm{T}}$$

where M^{T} is the transpose of the matrix D.

Proof. By the definition of matrix representation, we have

$$T^*(g_i) = \sum_{k=1}^{n} \beta_{ki} f_k \qquad (i = 1, \ldots, m)$$

and

$$T(v_j) = \sum_{i}^{m} \alpha_{ij} w_i \qquad (j = 1, \ldots, n).$$

It can be checked that

$$[T^*(g_i)](v_j) = (g_i \circ T)(v_j) = \alpha_{ij}$$

and

$$[T^*(g_i)](v_j) = \left(\sum_{k=1}^{n} \beta_{ki} f_k\right)(v_j) = \beta_{ji}$$

which implies $\alpha_{ij} = \beta_{ji}$.

□

Now we are able to prove the fact mentioned at the end of last section.

Proposition 9.5. *Let W be a subspace of a vector space V. Then $(V/W)^* \cong \mathrm{Ann}(W)$.*

Proof. For any $f \in \mathrm{Ann}(W)$, we have $W \leqslant \ker(f)$. By universal property 1 in 6.10, there exists $\bar{f}: V/W \longrightarrow \mathbb{F}$ such that

$$f = \bar{f} \circ \pi_W$$

where $\bar{f}(v + W) = f(v)$. Consider these two mappings

$$\Phi: \mathrm{Ann}(W) \longrightarrow (V/W)^*$$
$$\pi_W^*: (V/W)^* \longrightarrow V^*$$

where $\Phi(f) = \bar{f}$. We claim that π_W^* is an isomorphism.

It can be easily checked that $\pi_W^*((V/W)^*) \subseteq \operatorname{Ann}(W)$, by taking $h \in (V/W)^*$ and compute $[\pi_W^*(h)](w)$ for any $w \in W$.

It is also evident that π_W^* is injective since π_W is surjective.

The last step is to study the mapping $\pi_W^* \circ \Phi$. It is well-defined on $\operatorname{Ann}(W)$, so take $g \in \operatorname{Ann}(W)$ and it follows that

$$(\pi_W^* \circ \Phi)(g) = g.$$

Thus $\pi_W^* \circ \Phi$ is an identity mapping on $\operatorname{Ann}(W)$, and we can conclude that π_W^* is surjective.

□

Bibliography

[1] K.M. Hoffman and R.A. Kunze. *Linear Algebra*. Prentice-Hall Mathematics Series. N.J., Prentice-Hall, 1971.

[2] W. Rudin. *Principles of Mathematical Analysis*. International series in pure and applied mathematics. McGraw-Hill, 1976.